Journal of Reliability and Statistical Studies; ISSN (Print): 0974-8024, (Online):2229-5666
Vol. 5, Issue 2 (2012): 45-58

BAYESIAN INFERENCE FOR THE PARAMETER OF THE POWER DISTRIBUTION

Tanveer Kifayat[1] ,Muhammad Aslam[2] and Sajid Ali[3]
[1,2]Department of Statistics, Quaid-i-Azam University, Islamabad, Pakistan
E Mail: [2]aslamdqu@yahoo.com **;** : [1] tanveerqau@yahoo.com
[3]Department of Decesion Sciences, Bocconi University, Milan, Italy.
E Mail: [3]sajidali.qau@hotmail.com

(Received June 17, 2012)

Abstract

This study provides Bayesian analysis of the power model using two informative (gamma and Rayleigh) priors and two non-informative (Jeffreys and uniform) priors. The prior predictive distribution is used to elicit the values of the hyperparameters of the prior distribution. The priors are compared using Bayes point and interval estimates, posterior variances, coefficients of skewness and coefficients of kurtosis. Bayes factors and Bayes posterior risks are also used for the comparison of informative and non-informative priors.

Key Words: Bayes Factor, Bayes Posterior Risk, Elicitation, Hyperparameter,

1. Introduction

The power distribution is typically used as a subjective description of a population for which there is only limited sample data, and especially in cases where the relationship between variables is known but data is scare (possibly because of the high cost of collection).

Meniconi and Barry (1996) have explained many statistical distributions used in the assessment of semiconductor device and product reliability. But power function distribution is preferred over exponential, lognormal and Weibull among others because it exhibits a better fit for failure data and provides more appropriate information about reliability and hazard rates. Dallas (1976) has enlightened that if X follows power distribution, then X^{-1} follows the Pareto distribution. Saran and Pandey (2004) have put forward the concept of record values which are found in many situations of daily life as well as in many statistical applications. By using the order statistics they have obtained the best linear unbiased estimates of the parameter of the power function distribution in terms of *kth* upper record values. Chang (2007) presents characterizations of the power function distribution by independence of record values. Haq and Dey (2011) considered the Bayesian estimation of Erlang distribution using different informative and noninformative priors.

In this paper, the posterior distribution for the unknown parameter θ of the power distribution is derived using informative (gamma and Rayleigh) priors and non-informative (Jeffreys and uniform) priors. The prior predictive distribution under informative priors has been derived, which is used for the elicitation of hyperparameters. The paper is organized in the following sections. The power distribution is defined in Section 2 and it includes the derivation of the posterior distribution under non-informative priors. Section 3 comprises of derivation of posterior distribution under informative priors. Section 4 contains the detail of the method used for elicitation of hyperparameters. Section 5 provides the comparison of priors through posterior variances, coefficients of skewness, coefficients of kurtosis, Bayesian point estimates, credible intervals, Bayes factors for different hypotheses, Bayes estimators under different loss functions and Bayes posterior risks. Some concluding remarks are given in the last Section 6.

2. The Posterior Distribution Using Non-Informative Priors

Bayesian analysis is performed by combining the prior information $p(\theta)$ and the sample information $(x_1, x_2, ..., x_n)$ into what is called the posterior distribution of θ given $\mathbf{x} = x_1, x_2, ..., x_n$, from which all decisions and inferences are made. So $p(\theta|x)$ reflects the updated beliefs about θ after observing the sample $\mathbf{x} = x_1, x_2, ..., x_n$.

The posterior distributions using non-informative priors for the unknown parameter θ of the power distribution are derived below:

Let X be a random variable having the following p.d.f. $f(x)$ for a power distribution with unknown parameter θ:

$$f(x) = \theta x^{\theta-1},\ 0 < x < 1,\ 0 < \theta < \infty$$

The likelihood function of a random sample $(\mathbf{x} = x_1, x_2, ..., x_n)$ from the power distribution with unknown parameter θ is:

$$L(\theta, \mathbf{x}) = \theta^n \prod_{i=1}^{n} x_i^{\theta-1},\ 0 < x < 1;\ i = 1, 2, ..., n. \tag{2.1}$$

Prior probability distribution is a distribution of an uncertain quantity of θ, that would express one's uncertainty about θ before the data are taken into account. If there is no relevant prior information available then there are ways to derive a non-informative prior distribution. The parameters of prior distribution are called hyperparameters.

2.1 Posterior Distribution Using the Jeffreys Prior

A non-informative prior has been suggested by Jeffreys, which is frequently used in situation where one does not have much information about the parameters. This defines the density of the parameters proportional to the square root of the determinant of the Fisher information matrix, symbolically the Jeffreys prior of θ is:

$$p(\theta) \propto \sqrt{|I(\boldsymbol{\theta})|}$$

where $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_k)'$ be the vector of unknown parameters, and $I(\boldsymbol{\theta})$ is the $(k \times k)$ Fisher information matrix which is the logarithm of likelihood function $L(\theta)$ of parameter θ's and partially differentiating twice with respect to the parameters as given below:

$$I(\boldsymbol{\theta}) = -\boldsymbol{E}\left[\frac{\partial^2 lnL(\theta, \mathbf{x})}{\partial\theta_j \partial\theta_i}\right]$$

Here E denotes expectation on data and i and j stand for rows and columns of determinant respectively.

The Jeffreys prior for the parameter θ of the power distribution is derived which is:

$$p(\theta) \propto \theta^{-1}, \; 0<\theta<\infty \tag{2.2}$$

The posterior distribution of θ for the given data $(\mathbf{x} = x_1, x_2, ..., x_n)$ using equation (2.1) and (2.2) is:

$$p(\theta \mid \mathbf{x}) \propto p(\theta) L(\theta, \mathbf{x})$$

$$p(\theta \mid \mathbf{x}) \propto \theta^{-1} \theta^{n} \prod_{i=1}^{n} x_i^{\theta-1}$$

$$p(\theta \mid \mathbf{x}) \propto \theta^{n-1} e^{-\theta \sum_{i=1}^{n} \ln x^{-1}{}_{i}}$$

$$p(\theta \mid \mathbf{x}) \propto \theta^{\alpha_J} e^{-\beta_J \theta}, \; 0 < \theta < \infty \tag{2.3}$$

which is the density kernel of gamma distribution having parameters $\alpha_J = n$ and $\beta_J = \sum_{i=1}^{n} \ln x_i^{-1}$. So the posterior distribution of $\theta | \mathbf{x} \sim gamma(\alpha_J, \beta_J)$.

2.2 Posterior Distribution Using Uniform Prior

An obvious choice for the non-informative prior is the uniform distribution. Uniform priors are particularly easy to specify in the case of a parameter with bounded support. The uniform prior of θ is defined as:

$$p(\theta) \propto 1, 0 < \theta < \infty \tag{2.4}$$

The posterior distribution of parameter θ for the given data $(\mathbf{x} = x_1, x_2, ..., x_n)$ using (2.1) and (2.4) is:

$$p(\theta \mid \mathbf{x}) \propto \theta^{n} \prod_{i=1}^{n} x_i^{\theta-1}$$

$$p(\theta \mid \mathbf{x}) \propto \theta^{n+1-1} e^{-\theta \sum_{i=1}^{n} \ln x_i^{-1}}$$

$$p(\theta \mid \mathbf{x}) \propto \theta^{\alpha_U - 1} e^{-\beta_U \theta}, \; 0 < \theta < \infty \tag{2.5}$$

which is the density kernel of gamma distribution having parameters $\alpha_U = n+1$ and $\beta_U = \sum_{i=1}^{n} \ln x_i^{-1}$. So the posterior distribution of $\theta | \mathbf{x} \sim gamma(\alpha_J, \beta_J)$.

3. The Posterior Distribution Using Informative Priors

Here we use gamma and Rayleigh distribution as informative prior because they are compatible with the parameter θ of the power distribution. The posterior distributions using informative priors for the unknown parameter θ of the power distribution are derived below:

3.1 Posterior Distribution Using Gamma Prior

A way to guarantee that the posterior has an easily calculatable form is to specify a conjugate prior. Conjugacy is a joint property of the prior and the likelihood function that provides a posterior from the same distributional family as the prior. Gamma distribution is the conjugate prior of the power distribution. The gamma distribution is used as an informative prior with hyperparameters '*a*' and '*b*', having the following p.d.f.:

$$p(\theta) = \frac{b^a}{\Gamma a}\theta^{a-1}e^{-b\theta}\ ,\ 0<\theta<\infty,\ \ a,b>0 \qquad (3.1)$$

The posterior distribution of parameter θ for the given data $(\mathbf{x} = x_1, x_2, ..., x_n)$ using equations (2.1) and (3.1) is:

$$p(\theta \mid \mathbf{x}) \propto \frac{b^a}{\Gamma a}\theta^{a-1}e^{-b\theta}\theta^n \prod_{i=1}^{n} x_i^{\theta-1}$$

$$p(\theta \mid \mathbf{x}) \propto \theta^{a+n-1}\, e^{-\theta(b-\sum_{i=1}^{n}\ln x_i)}$$

$$p(\theta \mid \mathbf{x}) \propto \theta^{\alpha-1}\, e^{-\beta\theta}\ ,\ 0<\theta<\infty \qquad (3.2)$$

which is the density kernel of the gamma distribution having parameters $\alpha = a+n$ and $\beta = b + \sum_{i=1}^{n} \ln x_i^{-1}$. So the posterior distribution of $\theta \big| \mathbf{x} \sim gamma(\alpha, \beta)$.

3.2 Posterior Distribution Using Rayleigh Prior

Another informative prior is assumed to be the Rayleigh distribution with hyperparameter '*c*', which has the following p.d.f.:

$$p(\theta) = \frac{\theta\, e^{\frac{-\theta^2}{2c^2}}}{c^2},\ \ \theta>0,\ c>0 \qquad (3.3)$$

The posterior distribution of parameter θ for the given data $(\mathbf{x} = x_1, x_2, ..., x_n)$ using equation (2.1) and (3.3) is:

$$p(\theta \mid \mathbf{x}) \propto \theta^{n+1}\, e^{-\theta(\frac{\theta}{2c^2}+\sum_{i=1}^{n}\ln x_i^{-1})}$$

$$p(\theta \mid \mathbf{x}) = \frac{1}{k}\theta^{n+1}\, e^{-\theta(\frac{\theta}{2c^2}+\sum_{i=1}^{n}\ln x_i^{-1})}\ ,0<\theta<\infty \qquad (3.4)$$

$$\text{where } k = \int_0^\infty \theta^{n+1}\, e^{-\theta(\frac{\theta}{2c^2}+\sum_{i=1}^{n}\ln x_i^{-1})}\, d\theta$$

The posterior distribution is not in closed form but can be used numerically using package like SAS.

4. Elicitation of Hyperparameters

In the context of Bayesian statistical analysis, it arises most usually as a method for specifying the prior distribution for one or more unknown parameters of a statistical model.

Different methods of elicitation are proposed by Aslam(2003), here we choose the method of the confidence levels (C.L) of the prior predictive distribution to elicit the hyperparameters of the prior density.

For analysis we take the sample of 20 observations from Mendenhall and Hader (1958) mixture data recorded about times to failure for ARC-1 VHF communication transmitter receivers of a single commercial airline. Saleem et. al. (2010) have used the transformation $(x=\exp(-t))$ which yields a power distribution, so we also use this transformation. Following is the set of 20 observations:

152, 528, 424, 208, 536, 40, 8, 224, 112, 72, 72, 72, 112, 360, 120, 552, 104, 384, 464, 552.

$$\sum_{i=1}^{n} \ln x^{-1}{}_{i} = 5096 \text{ and } n = 20$$

4.1 The Prior Predictive Distribution

The prior predictive distribution or in other words the marginal distribution of an unobserved data value is the product of the prior for θ and the single variable p.d.f., integrating out this parameter. This makes intuitive sense as uncertainty in θ is averaged out to reveal a distribution for the data point. It is defined as:

$$p(y) = \int_{0}^{\infty} p(\theta) f(y;\theta)\, d\theta$$

here Y is the random variable of the model with unknown parameter θ.

$$f(y;\theta) = \theta\, y^{\theta-1}, \quad 0 < y < 1, \quad 0 < \theta < \infty. \tag{4.1}$$

4.2 Prior Predictive Distribution Using Gamma Prior

The prior predictive distribution or in other words the marginal distribution of an unobserved data value is the product of the prior for θ and the single variable p.d.f., integrating out this parameter. This makes intuitive sense as uncertainty in θ is averaged out to reveal a distribution for the data point.

The prior predictive distribution using gamma prior for a random variable Y combining equation (3.1) and (4.1) is:

$$p(y) = \frac{a\,b^{a}}{y\,(b + \ln y^{-1})^{a+1}}, \quad 0 < y < 1. \tag{4.2}$$

The equation of prior predictive distribution is used for the elicitation of the hyperparameters.

The two confidence levels for the prior predictive distribution may be elicited as 0.05 and 0.05 associated with the equation (4.2) of prior predictive distribution over the intervals $(e^{-155}$ to $e^{-130})$ and $(e^{-30}$ to $e^{-5})$.

$$0.05 = \int_{e^{(-155)}}^{e^{(-130)}} \frac{\mathrm{a\,b^{a}}\, dy}{\mathrm{y\,(b+\ln y^{-1})}^{(a+1)}}$$

$$0.05 = \int_{e^{(-30)}}^{e^{(-5)}} \frac{\mathrm{a\,b^{a}}\, dy}{\mathrm{y\,(b+\ln y^{-1})}^{(a+1)}}.$$

These two equations are solved simultaneously by applying 'PROC SYSLIN' in the SAS package for eliciting the hyperparameters '*a*' and '*b*'. In this way, we found that the values of the hyperparameters '*a*' and '*b*' are to be 2.4941 and 0.0936 respectively.

Now the posterior distribution of parameter θ using equation (3.2) is the gamma distribution with parameters $\alpha = 22.4941$ and $\beta = 5096.0936$.

4.3 Prior Predictive Distribution Using Rayleigh Prior

Further the prior predictive distribution using Rayleigh prior for a random variable *Y* combining equation (3.3) and (4.1) is:

$$p(y) = \frac{1}{y c^2} \int_0^\infty \theta^2 e^{-\theta(\frac{\theta}{2c^2} + \ln y^{-1})} d\theta, \; 0 < y < 1 \qquad (4.3)$$

This prior predictive distribution has not a closed form and it is used numerically for the elicitation of hyperparameter '*c*'.

The confidence level for the prior predictive distribution may be elicited as 0.05 associated with the equation (4.3) of prior predictive distribution over the interval $(0 \text{ to } e^{-8})$ considering the data set.

$$0.05 = \frac{1}{c^2} \int_0^{e^{-8}} \int_0^{20} \frac{\theta^2 e^{-\theta(\frac{\theta}{2c^2} + \ln y^{-1})}}{y} d\theta \, dy$$

We solve this equation by applying 'PROC SYSLIN' in the SAS package for eliciting the hyperparameter '*c*'. In this way, we find that the value of the hyperparameter '*a*' is to be 0.2291.

Using the sample information and the elicited hyperparameter, the posterior distribution of parameter θ using equation (3.4) is:

$$p(\theta|\mathbf{x}) = \frac{\theta^{21} e^{-\theta(\frac{\theta}{2(0.2291)^2} + 5096)}}{\int_0^\infty \theta^{21} e^{-\theta(\frac{\theta}{2(0.2291)^2} + 5096)} d\theta}, \; 0 < \theta < \infty.$$

The posterior distribution of parameter θ using equation (2.3) the Jeffreys prior is gamma distribution with parameters $\alpha_J = 20$ and $\beta_J = 5096$.

The posterior distribution of parameter θ using equation (2.5) uniform prior is also gamma distribution with parameters $\alpha_U = 21$ and $\beta_U = 5096$.

5. Bayes Estimators Under Different Loss Functions

This section focuses on the Bayes estimators and Bayes posterior risks under different loss functions and compares their results for the informative priors and non-informative priors. Loss function is a real valued function that explicitly provides a loss (penalty) for decision θ^* given θ is the true parameter value.

We use four loss functions. The derivations of Bayes estimates and Bayes posterior risk are given in appendix. Note that Bayes posterior risk $\rho_D(\theta^*)$ using absolute loss function is numerically solved. The values of the Bayes estimates under the mentioned loss functions are given in Table 5.0.

Loss Function $L(\theta,\theta^*)$	Bayes Estimator θ^*	Bayes Posterior Risk $\rho(\theta^*)$	Prior Distribution		Bayes Estimtor
L_A	$\frac{E(\theta^{-1})}{E(\theta^{-2})}$	$\rho_A(\theta^*)=1-\frac{\left(E(\theta^{-1})\right)^2}{E(\theta^{-2})}$	IP	Gamma	0.00402 (0.04652)
				Rayleigh	0.00392 (0.04762)
			NIP	Jeffreys	0.00353 (0.05632)
				Uniform	0.00373 (0.05000)
L_B	$\frac{1}{E(\theta^{-1})}$	$\rho_B(\theta^*)=E(\theta)-\frac{1}{E(\theta^{-1})}$	IP	Gamma	0.00422 (0.00020)
				Rayleigh	0.00412 (0.00020)
			NIP	Jeffreys	0.00373 (0.00020)
				Uniform	0.00392 (0.00020)
L_C	$E(\theta)$	$\rho_C(\theta^*)=E(\theta^2)-(E(\theta))^2$	IP	Gamma	0.00441 (8.661×10^{-7})
				Rayleigh	0.00432 (8.471×10^{-7})
			NIP	Jeffreys	0.00392 (7.701×10^{-7})
				Uniform	0.00412 (8.086×10^{-7})
L_D	*Median*	$\rho_D(\theta^*)=$ *Numerically Solved*	IP	Gamma	0.00435 (0.00074)
				Rayleigh	0.00425 (0.00073)
			NIP	Jeffreys	0.00386 (0.00070)
				Uniform	0.00405 (0.00071)

* IP stands for informative prior * NIP stands for non-informative prior

Table 5.0: Bayes Estimates Under Different Loss Functions

The Bayes estimators under loss functions $(L_A, L_B, L_C \text{ and } L_D)$ have minimum value for non-informative (Jeffreys) prior than the informative and non-informative (uniform) prior. The Bayes estimators under the loss function L_A have overall minimum value than the loss functions $(L_B, L_C \text{ and } L_D)$.

The values of Bayes posterior risk are given in the brackets. From Table 5.1 we conclude that the Bayes posterior risk $\rho_C(\theta^*)$ has overall smaller value than $(\rho_A(\theta^*), \rho_B(\theta^*)$ and $\rho_D(\theta^*))$ under the informative and non-informative priors.

5.1 Comparison of Posterior Variances

The posterior variances are given in Table 5.2:

Parameter	Posterior Variance			
	I P		N I P	
θ	Gamma	Rayleigh	Jeffreys	Uniform
	8.6615×10^{-7}	8.4710×10^{-7}	7.7011×10^{-7}	8.0862×10^{-7}

Table 5.1: Comparison of different Priors with respect to Posterior Variances

From table 5.1, it is observed that

Var (Jeffreys) < Var (Uniform) < Var (Rayleigh) < Var (Gamma)

From table 5.1 we infer that posterior variance using non-informative (Jeffreys) prior is less than the posterior variance of informative (gamma and Rayleigh) and non-informative (uniform) prior. So we conclude that the Jeffreys prior is more efficient prior among these four prior which we have used.

5.2 Credible Interval

Here we are interested in finding out the credible interval estimates. If a random variable X follows the power distribution then Meniconi and Barry (1996) mentioned that the credible interval for θ is given as:

$$p\left(\frac{\chi^2_{(\alpha,2n)}}{-2\sum_{i=1}^{n}\log x_i} \le \theta \le \frac{\chi^2_{(1-\alpha,2n)}}{-2\sum_{i=1}^{n}\log x_i}\right) = 1-\alpha$$

Prior Distributions		95% Credible Interval	99% Credible Interval
I P	Gamma	(0.00341, 0.00662)	(0.00291, 0.00747)
	Rayleigh	(0.00273, 0.00591)	(0.00198, 0.00666)
N I P	Jeffreys	(0.00260, 0.00547)	(0.00217, 0.00625)
	Uniform	(0.00341, 0.00662)	(0.00292, 0.00747)

Table 5.2: Credible Intervals

The Table 5.2 compares credible intervals, we find out that 95% credible interval are narrow than 99% interval for informative and non-informative priors. Furthermore, we conclude that 99% credible interval for informative (Rayleigh) is greater than informative (gamma) prior. We observe that for non-informative priors, 99% credible interval considering Jeffreys prior is narrower than uniform prior. By examining 95% credible interval for informative (Rayleigh) is narrower than informative (gamma) prior and for non-informative, uniform prior is wider than Jeffreys prior.

5.3 Coefficient of Skewness and Coefficient of Kurtosis

As the posterior distribution is skewed so we calculate the coefficients of skewness and kurtosis of the posterior distributions and given in Table 5.3.

Prior Distribution		Moments about Mean				C.O.S	C.O.K
		μ_1	μ_2	μ_3	μ_4	γ_1	γ_2
I P	Gamma	0	8.6615×10^{-7}	3.3993×10^{-10}	2.4508×10^{-12}	0.42169	3.26674
	Rayleigh	0	8.4711×10^{-7}	3.2916×10^{-10}	2.3795×10^{-12}	0.42162	3.26668
N I P	Jeffreys	0	7.7014×10^{-6}	3.0225×10^{-10}	1.9573×10^{-12}	0.44721	3.30000
	Uniform	0	8.0865×10^{-6}	3.1737×10^{-10}	2.1486×10^{-12}	0.43644	3.28571

* C.O.S stands for Coefficient of Skewness, * C.O.K stands for Coefficient of Kurtosis

Table 5.3: Coefficients of Skewness and Kurtosis

The Table 5.3 shows that the coefficients of skewness for both informative and non-informative priors are greater than zero, so the posterior distributions for all these priors are positively skewed. As the coefficient of kurtosis is greater than 3 for both the informative and non-informative priors then they are leptokurtic distribution having a more acute peak around the mean.

5.4 Bayes Factor for Bayesian Hypotheses Testing

Bayesian hypothesis testing describes as the evidence of the quality of one model specification over another. This section contains the testing of parameter θ considering different null and alternative hypotheses using informative and non-informative priors. The arbitrary decision thresholds for these hypotheses are based on Jeffreys (1961) typology for comparing model H_0 and H_1.

Null Hypotheses	Alternative Hypotheses	Prior Distribution		Posterior Probability		Bayes Factor
H_0	H_1			$P(H_0)$	$P(H_1)$	B
$\theta \le 0.0035$	$\theta > 0.0035$	I P	Gamma	0.16163	0.83837	0.19279
			Rayleigh	0.18142	0.81858	0.22163
		N I P	Jeffreys	0.33461	0.66539	0.50287
			Uniform	0.25631	0.74369	0.34465
$\theta \le 0.0040$	$\theta > 0.0040$	I P	Gamma	0.34878	0.65122	0.53558
			Rayleigh	0.37816	0.62184	0.60813
		N I P	Jeffreys	0.56349	0.43651	1.29090
			Uniform	0.47498	0.52502	0.90468
$\theta \le 0.0045$	$\theta > 0.0045$	I P	Gamma	0.56437	0.43563	1.29553
			Rayleigh	0.59505	0.40495	1.46944
		N I P	Jeffreys	0.75798	0.24202	3.13189

			Uniform	0.68496	0.31504	2.17420
$\theta \le 0.0050$	$\theta > 0.0050$	I P	Gamma	0.74965	0.25035	2.99441
			Rayleigh	0.77393	0.22607	3.42341
		N I P	Jeffreys	0.88524	0.11476	7.71383
			Uniform	0.83824	0.16176	5.18199

Table 5.4 Posterior Probabilities and Bayes Factor for Differen t Hypotheses

In Table 5.4 we calculate Bayes Factor (B), central notion is that prior and posterior information should be combined in a ratio that provides evidence of one model specification over another. From table 5.4 we observed that for hypothesis $H_0 : \theta \le 0.0035$ Vs $H_1 : \theta > 0.0035$ we have substantial evidence against H_0 for informative priors and have minimal evidence against H_0 for non-informative priors. Considering $H_0 : \theta \le 0.0040$ Vs $H_1 : \theta > 0.0040$ we have minimal evidence against H_0 for informative and non-informative (uniform) priors and supported H_0 for non-informative (Jeffreys) prior. Similarly for hypotheses $H_0 : \theta \le 0.0045$ Vs $H_1 : \theta > 0.0045$ and $H_0 : \theta \le 0.0050$ Vs $H_1 : \theta > 0.0050$ we have supported H_0 for both informative and non-informative priors.

6. Conclusion

By carrying out this study we enlighten the Bayesian analysis of the power model using informative (gamma and Rayleigh) and non-informative (Jeffreys and uniform) priors. From Table 5.0, we observe that the posterior risk is smaller for non-informative than the informative priors. The result of the Table 5.2 shows that 95% credible interval are narrow than 99% interval for informative and non-informative priors. The variance of Jeffreys prior is more efficient among all the priors given in Table 5.1. The posterior mean is greater than posterior median, so we conclude that our posterior distributions are skewed for informative and non-informative priors. Coefficients of skewness in Table 5.3 are positive, so the posterior distributions using all these priors are positively skewed. Coefficients of kurtosis using informative and non-informative priors are greater than 3, and then they are leptokurtic distributions having a more acute peak around the mean. The testing of parameter θ, in Table 5.4 considering different null and alternative hypotheses using informative and non-informative priors, when the Bayes factor B > 1 in such a case we have supported H_0. The Bayes estimators under the different loss functions have minimum value for non-informative (Jeffreys) prior than the informative and non-informative (uniform) priors.

References

1. Aslam, M. (2003). An Application of Prior Predictive Distribution to Elicit the Prior Density, Journal of Statistical Theory and Applications, 2 (1), p. 70-83.
2. Berger, J. O. (1985). Statistical Decision Theory and Bayesian, 2nd ed. New York: Springer-Verlag,
3. Berger, J.O. (1985). Statistical Decision Theory and Bayesian, 2nd Edn. Springer-Verlag, New York.
4. Bolstad, W.M. (2004). Introduction to Bayesian Statistics, John Wiley & Sons.

5. Chang, S.K. (2007). Characterizations of the Power Function Distribution by the Independence of Record Values, Journal of the Chungcheong Mathematical Society, 20 (2), p. 139-146.
6. Dallas, A.C. (1976). Characterizing the Pareto and Power Distributions, Ann. Inst. Statist. Math., 28, p. 491-497.
7. Garthwaite, P.H., Kadane, J.B. and O'Hagan, A. (2005). Statistical Methods for Eliciting Probability Distributions, Journal of the American Statistical Association, 100 (470), p. 680-701.
8. Gill, J. (2008). Bayesian Methods: A Social and Behavioural Sciences Approach, 2nd ed. New York: Chapman and Hall.
9. Haq, A. and Dey, S. (2011). Bayesian estimation of the Erlang distribution under different prior distributions, Journal of Reliability and Statistical Studies, 4 (1), p. 1-30.
10. Jeffreys, H. (1946). An Invariant Form for the Prior Probability in Estimation Problems, Journal Proceeding of the Royal Society Of London, Series A, 186 (1007), p. 453-461.
11. Jeffreys, H. (1961).Theory of Probability, 3rd ed. Oxford Univ. Press.
12. Mendenhall, W., and Hader, R.J. (1958). Estimation of Parameters of Mixed Exponentially Distributed Failure Time Distributions from Censored Life Test Data, Biometrika, 45 (3), p. 504-520.
13. Meniconi, M. and Barry, D.M. (1996). The Power Function Distribution: A Useful and Simple Distribution to Assess Electrical Component Reliability, Journal Microelectron Reliab. , 36 (9), p. 1207-1212.
14. Saleem, M., Aslam, M., and Economou, P. (2010). On the Bayesian Analysis of the Mixture of Power Function Distribution Using the Complete and the Censored 'Sample', Journal of Applied Statistics, 37 (1), p. 25-40.
15. Saran, J. and Pandey, A. (2004). Estimation of Parameters of a Power Function Distribution and Its Characterization by K^{th} Record Values. Journal Statistica, 14 (3), p. 523-536.

Appendix

We have used four loss functions for estimation. The derivations of the Bayes estimators are given in this appendix.

1. Quadratic Loss Function

The quadratic loss function is given as:

$$L_A = L_A(\theta,\theta^*) = \left(1-\frac{\theta^*}{\theta}\right)^2$$

The posterior risk function is:

$$\rho_A(\theta^*) = E\left(1-\frac{\theta^*}{\theta}\right)^2$$

$$\rho_A(\theta^*) = 1-2\theta^* E(\theta^{-1}) + \theta^{*2} E(\theta^{-2}) \qquad \text{(A)}$$

Differentiating equation (A) w.r.t θ^* we get

$$\frac{\partial \rho_A(\theta^*)}{\partial \theta^*} = -2E(\theta^{-1}) + 2\theta^* E(\theta^{-2})$$

For minimizing, $\frac{\partial \rho_A(\theta^*)}{\partial \theta^*} = 0$ as 2nd derivative is positive.

$$\theta^* = \frac{E(\theta^{-1})}{E(\theta^{-2})} \qquad \text{(B)}$$

which is Bayes estimator under quadratic loss function.

2. Weighted Loss Function

The weighted loss function is given as:

$$L_B = L_B(\theta,\theta^*) = \frac{(\theta-\theta^*)^2}{\theta}$$

The Bayes estimator θ^* is derived by the rule of calculus as applied in quadratic loss function. So the Bayes estimator is:

$$\theta^* = \frac{1}{E(\theta^{-1})}$$

Squared Error Loss Function

The squared error loss function is given as:

$$L_C = L_C(\theta,\theta^*) = (\theta-\theta^*)^2$$

The Bayes estimator using above loss function is:

$$\theta^* = E(\theta)$$

which is the posterior mean.

3. Absolute Loss Function

The absolute loss function is given as:

$$L_D = L_D(\theta,\theta^*) = \left|\theta-\theta^*\right|$$

The posterior risk is:

$$\rho_D(\theta^*) = E_{\theta|\mathbf{x}} L_D(\theta,\theta^*) = E_{\theta|\mathbf{x}} \left|\theta-\theta^*\right|$$

$$\rho_D(\theta^*) = \int_{-\infty}^{\infty} \left|\theta-\theta^*\right| p(\theta|\mathbf{x})d\theta$$

$$\rho_D(\theta^*) = \int_{-\infty}^{\theta^*} -(\theta-\theta^*)p(\theta|\mathbf{x})d\theta \ + \int_{\theta^*}^{\infty} \theta-\theta^*)p(\theta|\mathbf{x})d\theta \qquad \text{(C)}$$

$$\rho_D(\theta^*) = 2\theta^* F(\theta^*|\mathbf{x}) - \theta^* - 2\int_{-\infty}^{\theta^*} \theta p(\theta|\mathbf{x})d\theta \ +1$$

Differentiating w.r.t θ^* we get

For minimizing $\dfrac{\partial \rho_D(\theta^*)}{\partial \theta^*} = 2F(\theta^*|\mathbf{x}) + 2\theta^* p(\theta^*|\mathbf{x}) \ -1 - 2\theta^* p(\theta^*|\mathbf{x})$

$$\frac{\partial \rho_D(\theta^*)}{\partial \theta^*} = 0 \qquad \Rightarrow 2F(\theta^*|\mathbf{x}) \ -1 = 0$$

$$\theta^* = F(\theta^*|\mathbf{x}) \ = \frac{1}{2}$$

According to definition of median in terms of distribution function; $\theta \mid \mathbf{x}$ is a posterior median which is Bayes estimator under absolute loss function.

Bayes Posterior Risk

This section contains the derivation of Bayes posterior risk using different loss function.

1. The Bayes posterior risk using quadratic loss function is:

$$\rho_A = E_{\theta|\mathbf{x}} L_A(\theta,\theta^*)$$

Using equation (A) and put the value of equation (B) we get

$$\rho_A = \left\{ 1 - 2\frac{E(\theta^{-1})}{E(\theta^{-2})} E(\theta^{-1}) + \left(\frac{E(\theta^{-1})}{E(\theta^{-2})} \right)^2 E(\theta^{-2}) \right\}$$

$$\rho_A = 1 - 2\frac{\left(E(\theta^{-1})\right)^2}{E(\theta^{-2})} + \frac{\left(E(\theta^{-1})\right)^2}{E(\theta^{-2})}$$

$$\rho_A = 1 - \frac{\left\{E(\theta^{-1})\right\}^2}{E(\theta^{-2})}.$$

2. The Bayes posterior risk using weighted loss function is:

$$\rho_B(\theta^*) = E(\theta) - \frac{1}{E(\theta^{-1})}.$$

The Bayes posterior risk using squared error loss function is:

$$\rho_C(\theta^*) = E(\theta^2) - \{E(\theta)\}^2 .$$

which is the posterior variance of parameter θ and it is the Bayes posterior risk under the SLF.

3. The Bayes posterior risk $\rho_D(\theta^*)$ using absolute loss function is numerically solved.

   As $\theta^* = \text{Median}$ put in equation (C) to find the Bayes posterior risk :

$$\rho_D(\theta^*) = \int_{-\infty}^{\theta^*} -(\theta-\theta^*)p(\theta|\mathbf{x})d\theta \; + \int_{\theta^*}^{\infty} \theta-\theta^*)p(\theta|\mathbf{x})d\theta$$

   This equation is numerically solved.